

0590
0111

2 OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/014,269

DATE: 01/02/2002
TIME: 14:45:24

Input Set : A:\um06645.txt
Output Set: N:\CRF3\01022002\J014269.raw

ENTERED

3 <110> APPLICANT: Nunez, Gabriel
4 Inohara, Naohiro
5 Ogur, Yasunori
7 <120> TITLE OF INVENTION: NOD2 Nucleic Acids and Proteins
9 <130> FILE REFERENCE: UM-06645
C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/014,269
C--> 11 <141> CURRENT FILING DATE: 2001-10-26
11 <160> NUMBER OF SEQ ID NOS: 52
13 <170> SOFTWARE: PatentIn version 3.1
15 <210> SEQ ID NO: 1
16 <211> LENGTH: 4485
17 <212> TYPE: DNA
18 <213> ORGANISM: Homo sapiens
20 <400> SEQUENCE: 1
21 gtagacagat ccaggctcac cagtcctgtg ccactgggct tttggcggtc tg
23 ctaccgcag atgccatgcc tgctcccca gcctaattggg ctttgatggg gg
25 ggttcagcct ctcacgatga ggagaaaaaaga gcaagtgtcc tcctcgacaa tt
27 tgtgaaatgt gctcgcagga ggctttcag gcacagagga gccagctggt cg
29 gtctcagggt ccctggaagg cttcgagagt gtccctggact ggctgctgtc ct
31 ctctcctggg aggactacga gggcttccac ctctctggcc agcctctctc cc
33 aggccgccttc tgacaccgt ctggaataaag gtaattttggg cctgtcagaa gc
35 gtcgccaag aagcccagggc cgacagccag tcccccaagc tgcattggctg ct
37 cactcgctcc acccagcccg agacctgcag agtcaccggc cagccattgt ca
39 cacagccatg tgagaaacat gctggacactg gcatggggc ggggtttcgt ca
41 gaatgtgtat aaatcagggtt gccgatcttc acaccgtccc agagggcaag aa
43 gatcttgcca cggtaaaagc gaatggattt gtcgccttcc ttctacaaca tg
45 ttaccagtcc cattggccct gcctttggaa gtcgcacat gcaagaagta ta
47 ctgaggacca cgggtctgc tcagtctcgc ttcttcagta cctatgatgg ag
49 ctctgcctgg aggacatata cacagagaat gtccctggagg tctgggcaga tg
51 gctggacccc cgccagaagag cccagccacc ctgggcctgg aggagcttt ca
53 ggccacctca atgacgatgc ggacactgtg ctgggtgtgg gtgaggcggg ca
55 agcacgtcc tgacgggtc gcacttgcgt tgggctgcag ggcaagactt cc
57 ctctttgtct tcccatcag ctgcggcag ctgcagtgc tggccaaacc ac
59 cggactctac tctttgagca ctgcgttgg cctgatgttgc tcaagaaga ca
61 ttactccttgc accaccctga ccgtgttctg ttaacctttt atggctttga cg
63 ttcaggttca cgatcgtga acgccactgc tccccgaccg accccacctc tg
65 ctgctcttca accttctgca gggcaacctg ctgaagaatgg cccgcaagggt gg
67 cgtccggcccg ctgtgtcgcc gttcctcagg aagtacatcc gcaccggatg ca
69 ggcttccttg aacagggcat cgagctgtac ctgaggaagc gccatcatga gc
71 gcgacccgc tcacccgcct gctccaagag acctcagccc tgcacgggtt gt
73 cctgtcttct catggatgggt gtccaaatgc caccaggaac tttgtgcga gg
75 tccccaaaga ccactacaga tatgtacctg ctgattctgc agcattttct gc
77 acccccccac actcagcttc ccaaggctg ggacccagtc ttctcgggg cc
79 accctcctgc acctgggcag actggctctg tggggcctgg gcatgtgtc ct
81 tcagccacgc agctccaggc agcacaggc acggctgtatg acattttct tg
83 gtgcgtgcca aaggtgtcgt gccagggagt acggcgcccc tggaaattcct tc
85 ttccagtgtct tctttgcgc gttctacctg gcactcagtg ctgatgtgcc ac

RAW SEQUENCE LISTING DATE: 01/02/2002
PATENT APPLICATION: US/10/014,269 TIME: 14:45:24

Input Set : A:\um06645.txt
Output Set: N:\CRF3\01022002\J014269.raw

87 ctcagacacc tcttcaattg tggcaggccca ggcaactcac caatggccag gtcctgtccc
 89 acgatgtca tccaggcctc ggagggaaag gacagcagcg tggcagctt gtcgcagaag 2100
 91 gcccggcc acaaacctca gatcacagca gccttcctgg cagggctgtt gtcccgaggag 2160
 93 cactggggcc tgctggctga gtgccagaca tctgagaagg ccctgctccg ggcggccaggcc 2220
 95 tgtgcccgtt ggtgtctggc cccgagccctc cgcaagact tccactccat cccgccaagct 2280
 97 gcaccgggtg agggcaagag cgtgcattcc atgccccgggt tcatactggct catccggagc 2340
 99 ctgtacgaga tgcaggagga gcggtctggct cggaaggctg cacgtggcct gaatgttggg 2400
 101 cacctaagt tgacattttg cagtgtggc cccactgagt gtgctgcctt ggccttgtg 2460
 103 ctgcagcacc tccggccggcc cgtggcccttg cagctggact acaaactctgt ggggtgacatt 2520
 105 ggcgtggagc agctgtcgcc ttgccttgggt gtctgcaagg ctctgttattt ggcgcataac 2580
 107 aatatcttag accgaggcat ctgcaagctc attgaatgtt ctcttactg cgagcaattt 2640
 109 cagaagttag ctctattcaa caacaattt actgacggct gtgcacactc catggctaag 2700
 111 ctccattgtcat gcaggcagaa ctcttggca ttgaggctgg ggaataacta catcaactgcc 2760
 113 gccccggcc aagtgtggc cgaggggctc cgaggaaca cctccttgc gttccctggg 2820
 115 ttctggggca acagagtggg tgacggggg gcccaggccc tggctgaagc cttgggtgat 2880
 117 caccagagct tgagggtggct cagcctgggtt gggaaacaaca ttggcagtgt ggggtcccaa 2940
 119 gccttggcac ttagtgcgtt aaaaacgttc atgttagaaag aactctgcct ggaggagaac 3000
 121 catctccagg atgaagggtt atgttctctc gcagaaggac tgaagaaaaaa ttcaagttt 3060
 123 aaaatcctga agttgtccaa taactgcattt acctacccat gggcagaagc cctcctgcag 3120
 125 gccttggaaa ggaatgacac catcctggaa gtctgctcc gagggaaacac tttctctcta 3180
 127 gaggaggtt acaagctcg ctgcaggggac accagactt tgctttaag tctccgggag 3240
 129 gatgttcgtc tcagtttgg tttgagcggg ctgtgagttt gggcccccaga ggctgggtga 3300
 131 catgtgttgg cagcctcttc aaaatgagcc ctgtcttgc taaggctgaa cttgttttct 3360
 133 gggAACACCA taggtcacct ttattctggc agaggaggga gcatcagtgc cttccaggat 3420
 135 agacttttcc caagccatctt ttgcatttgc acttcttccc aagattcaat cccaggatgt 3480
 137 acaaggacag cccctctcc atagtatggg actggcctct gctgatcctc ccaggcttcc 3540
 139 gtgtgggtca gtggggccca tgatgtgtc tgtaactga gtgcctttt gtggagaggc 3600
 141 ccggcctctc acaaaaagacc ctttaccact gctctgtatgc agaggagttt acagaacaca 3660
 143 taattcagga agcagcttcc cccatgtctc gactcatcca tccaggccat tcccccgtctc 3720
 145 tggttcttcc ctttcccttgc gactcttgc cacgctcctt cctctgaggc taaaattcag 3780
 147 aatatttagtgc acctcagtt tgatatttca cttagcactt ccccaaccct ggcacccagg 3840
 149 gtgggaaggg ctacacccat gcttgccttc ctttccgggt ttttggaaat ttttggaaagg 3900
 151 ggacacgtga cagccgtttt ttcccccaaga cattcttaggt ttgcaagaaa aatatgacca 3960
 153 cactccagct gggatcacat gtggactttt atttccagttt aaatcagttt ctcttcagtt 4020
 155 aaggcccttgg aaacagctcg actttaaaaaa gtcacaaatg cagttttaaa aaattaatct 4080
 157 gggccagaat ttcaaacggc ctttcccttc ttctgggttgc tgcctgtgaa ctgaactctg 4140
 159 acaacagact tctgaaatag accccacaaga ggcaggccat tttcatttgc gccagaatgc 4200
 161 ttttaggatgtt acagttatgg attgaaatgtt tacagaaaaaaa aatatttgc ctttcccttc 4260
 163 aagcaaatgtt ctttcccttc ttttggatgtt gatgtatgtt gaagcctttt taaattgtca 4320
 165 gatgtgtgc aaatgttattt attttaaaca ttatgtgtt tgaaaactgg ttaatattt 4380
 167 taggtcactt tggtttactg tottaagttt atactttat agacaacatg gccgtgaact 4440
 169 ttatgtgtatc aataatcaga gggggataaaa ctgttggatc aaaaac 4485
 172 <210> SEQ ID NO: 2
 173 <211> LENGTH: 1040
 174 <212> TYPE: PRT
 175 <213> ORGANISM: Homo sapiens
 177 <400> SEQUENCE: 2
 179 Met Gly Glu Glu Gly Gly Ser Ala Ser His Asp Glu Glu Glu Arg Ala
 180 1 5 10 15

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/014,269

DATE: 01/02/2002

TIME: 14:45:24

Input Set : A:\um06645.txt

Output Set: N:\CRF3\01022002\J014269.raw

183 Ser Val Leu Leu Gly His Ser Pro Gly Cys Glu Met Cys Ser Gln Glu
 184 20 25 30
 187 Ala Phe Gln Ala Gln Arg Ser Gln Leu Val Glu Leu Leu Val Ser Gly
 188 35 40 45
 191 Ser Leu Glu Gly Phe Glu Ser Val Leu Asp Trp Leu Leu Ser Trp Glu
 192 50 55 60
 195 Val Leu Ser Trp Glu Asp Tyr Glu Gly Phe His Leu Leu Gly Gln Pro
 196 65 70 75 80
 199 Leu Ser His Leu Ala Arg Arg Leu Leu Asp Thr Val Trp Asn Lys Gly
 200 85 90 95
 203 Thr Trp Ala Cys Gln Lys Leu Ile Ala Ala Gln Glu Ala Gln Ala
 204 100 105 110
 207 Asp Ser Gln Ser Pro Lys Leu His Gly Cys Trp Asp Pro His Ser Leu
 208 115 120 125
 211 His Pro Ala Arg Asp Leu Gln Ser His Arg Pro Ala Ile Val Arg Arg
 212 130 135 140
 215 Leu His Ser His Val Glu Asn Met Leu Asp Leu Ala Trp Glu Arg Gly
 216 145 150 155 160
 219 Phe Val Ser Gln Tyr Glu Cys Asp Glu Ile Arg Leu Pro Ile Phe Thr
 220 165 170 175
 223 Pro Ser Gln Arg Ala Arg Arg Leu Leu Asp Leu Ala Thr Val Lys Ala
 224 180 185 190
 227 Asn Gly Leu Ala Ala Phe Leu Leu Gln His Val Gln Glu Leu Pro Val
 228 195 200 205
 231 Pro Leu Ala Leu Pro Leu Glu Ala Ala Thr Cys Lys Lys Tyr Met Ala
 232 210 215 220
 235 Lys Leu Arg Thr Thr Val Ser Ala Gln Ser Arg Phe Leu Ser Thr Tyr
 236 225 230 235 240
 239 Asp Gly Ala Glu Thr Leu Cys Leu Glu Asp Ile Tyr Thr Glu Asn Val
 240 245 250 255
 243 Leu Glu Val Trp Ala Asp Val Gly Met Ala Gly Pro Pro Gln Lys Ser
 244 260 265 270
 247 Pro Ala Thr Leu Gly Leu Glu Leu Phe Ser Thr Pro Gly His Leu
 248 275 280 285
 251 Asn Asp Asp Ala Asp Thr Val Leu Val Val Gly Glu Ala Gly Ser Gly
 252 290 295 300
 255 Lys Ser Thr Leu Leu Gln Arg Leu His Leu Leu Trp Ala Ala Gly Gln
 256 305 310 315 320
 259 Asp Phe Gln Glu Phe Leu Phe Val Phe Pro Phe Ser Cys Arg Gln Leu
 260 325 330 335
 263 Gln Cys Met Ala Lys Pro Leu Ser Val Arg Thr Leu Leu Phe Glu His
 264 340 345 350
 267 Cys Cys Trp Pro Asp Val Gly Gln Glu Asp Ile Phe Gln Leu Leu Leu
 268 355 360 365
 271 Asp His Pro Asp Arg Val Leu Leu Thr Phe Asp Gly Phe Asp Glu Phe
 272 370 375 380
 275 Lys Phe Arg Phe Thr Asp Arg Glu Arg His Cys Ser Pro Thr Asp Pro
 276 385 390 395 400
 279 Thr Ser Val Gln Thr Leu Leu Phe Asn Leu Leu Gln Gly Asn Leu Leu

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/014,269

DATE: 01/02/2002
TIME: 14:45:24

Input Set : A:\um06645.txt
Output Set: N:\CRF3\01022002\J014269.raw

280	405	410	415	
283	Lys Asn Ala Arg Lys Val Val Thr Ser Arg Pro Ala Ala Val Ser Ala			
284	420	425	430	
287	Phe Leu Arg Lys Tyr Ile Arg Thr Glu Phe Asn Leu Lys Gly Phe Ser			
288	435	440	445	
291	Glu Gln Gly Ile Glu Leu Tyr Leu Arg Lys Arg His His Glu Pro Gly			
292	450	455	460	
295	Val Ala Asp Arg Leu Ile Arg Leu Leu Gln Glu Thr Ser Ala Leu His			
296	465	470	475	480
299	Gly Leu Cys His Leu Pro Val Phe Ser Trp Met Val Ser Lys Cys His			
300	485	490	495	
303	Gln Glu Leu Leu Gln Glu Gly Gly Ser Pro Lys Thr Thr Asp			
304	500	505	510	
307	Met Tyr Leu Leu Ile Leu Gln His Phe Leu Leu His Ala Thr Pro Pro			
308	515	520	525	
311	Asp Ser Ala Ser Gln Gly Leu Gly Pro Ser Leu Leu Arg Gly Arg Leu			
312	530	535	540	
315	Pro Thr Leu Leu His Leu Gly Arg Leu Ala Leu Trp Gly Leu Gly Met			
316	545	550	555	560
319	Cys Cys Tyr Val Phe Ser Ala Gln Gln Leu Gln Ala Ala Gln Val Ser			
320	565	570	575	
323	Pro Asp Asp Ile Ser Leu Gly Phe Leu Val Arg Ala Lys Gly Val Val			
324	580	585	590	
327	Pro Gly Ser Thr Ala Pro Leu Glu Phe Leu His Ile Thr Phe Gln Cys			
328	595	600	605	
331	Phe Phe Ala Ala Phe Tyr Leu Ala Leu Ser Ala Asp Val Pro Pro Ala			
332	610	615	620	
335	Leu Leu Arg His Leu Phe Asn Cys Gly Arg Pro Gly Asn Ser Pro Met			
336	625	630	635	640
339	Ala Arg Leu Leu Pro Thr Met Cys Ile Gln Ala Ser Glu Gly Lys Asp			
340	645	650	655	
343	Ser Ser Val Ala Ala Leu Leu Gln Lys Ala Glu Pro His Asn Leu Gln			
344	660	665	670	
347	Ile Thr Ala Ala Phe Leu Ala Gly Leu Leu Ser Arg Glu His Trp Gly			
348	675	680	685	
351	Leu Leu Ala Glu Cys Gln Thr Ser Glu Lys Ala Leu Leu Arg Arg Gln			
352	690	695	700	
355	Ala Cys Ala Arg Trp Cys Leu Ala Arg Ser Leu Arg Lys His Phe His			
356	705	710	715	720
359	Ser Ile Pro Pro Ala Ala Pro Gly Glu Ala Lys Ser Val His Ala Met			
360	725	730	735	
363	Pro Gly Phe Ile Trp Leu Ile Arg Ser Leu Tyr Glu Met Gln Glu Glu			
364	740	745	750	
367	Arg Leu Ala Arg Lys Ala Ala Arg Gly Leu Asn Val Gly His Leu Lys			
368	755	760	765	
371	Leu Thr Phe Cys Ser Val Gly Pro Thr Glu Cys Ala Ala Leu Ala Phe			
372	770	775	780	
375	Val Leu Gln His Leu Arg Arg Pro Val Ala Leu Gln Leu Asp Tyr Asn			
376	785	790	795	800

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/014,269

DATE: 01/02/2002
TIME: 14:45:24

Input Set : A:\um06645.txt
Output Set: N:\CRF3\01022002\J014269.raw

379 Ser Val Gly Asp Ile Gly Val Glu Gln Leu Leu Pro Cys Leu Gly Val
380 805 810 815
383 Cys Lys Ala Leu Tyr Leu Arg Asp Asn Asn Ile Ser Asp Arg Gly Ile
384 820 825 830
387 Cys Lys Leu Ile Glu Cys Ala Leu His Cys Glu Gln Leu Gln Lys Leu
388 835 840 845
391 Ala Leu Phe Asn Asn Lys Leu Thr Asp Gly Cys Ala His Ser Met Ala
392 850 855 860
395 Lys Leu Leu Ala Cys Arg Gln Asn Phe Leu Ala Leu Arg Leu Gly Asn
396 865 870 875 880
399 Asn Tyr Ile Thr Ala Ala Gly Ala Gln Val Leu Ala Glu Gly Leu Arg
400 885 890 895
403 Gly Asn Thr Ser Leu Gln Phe Leu Gly Phe Trp Gly Asn Arg Val Gly
404 900 905 910
407 Asp Glu Gly Ala Gln Ala Leu Ala Glu Ala Leu Gly Asp His Gln Ser
408 915 920 925
411 Leu Arg Trp Leu Ser Leu Val Gly Asn Asn Ile Gly Ser Val Gly Ala
412 930 935 940
415 Gln Ala Leu Ala Leu Met Leu Ala Lys Asn Val Met Leu Glu Glu Leu
416 945 950 955 960
419 Cys Leu Glu Glu Asn His Leu Gln Asp Glu Gly Val Cys Ser Leu Ala
420 965 970 975
423 Glu Gly Leu Lys Lys Asn Ser Ser Leu Lys Ile Leu Lys Leu Ser Asn
424 980 985 990
427 Asn Cys Ile Thr Tyr Leu Gly Ala Glu Ala Leu Leu Gln Ala Leu Glu
428 995 1000 1005
431 Arg Asn Asp Thr Ile Leu Glu Val Trp Leu Arg Gly Asn Thr Phe
432 1010 1015 1020
435 Ser Leu Glu Glu Val Asp Lys Leu Gly Cys Arg Asp Thr Arg Leu
436 1025 1030 1035
439 Leu Leu
440 1040
443 <210> SEQ ID NO: 3
444 <211> LENGTH: 1013
445 <212> TYPE: PRT
446 <213> ORGANISM: Homo sapiens
448 <400> SEQUENCE: 3
450 Met Cys Ser Gln Glu Ala Phe Gln Ala Gln Arg Ser Gln Leu Val Glu
451 1 5 10 15
454 Leu Leu Val Ser Gly Ser Leu Glu Gly Phe Glu Ser Val Leu Asp Trp
455 20 25 30
458 Leu Leu Ser Trp Glu Val Leu Ser Trp Glu Asp Tyr Glu Gly Phe His
459 35 40 45
462 Leu Leu Gly Gln Pro Leu Ser His Leu Ala Arg Arg Leu Leu Asp Thr
463 50 55 60
466 Val Trp Asn Lys Gly Thr Trp Ala Cys Gln Lys Leu Ile Ala Ala Ala
467 65 70 75 80
470 Gln Glu Ala Gln Ala Asp Ser Gln Ser Pro Lys Leu His Gly Cys Trp
471 85 90 95

VERIFICATION SUMMARY
PATENT APPLICATION: US/10/014,269

DATE: 01/02/2002
TIME: 14:45:25

Input Set : A:\um06645.txt
Output Set: N:\CRF3\01022002\J014269.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application No
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date